Deployment Complexity Versus Performance Efficiency in Mobile Multicast

Anargyros Garyfalos
British Telecommunications Place
Intelligent Systems Lab
Adastral Park
Ipswich IP5 3RE, UK
anargyros.garyfalos@bt.com

Kevin C. Almeroth
Dept of Computer Science
University of California
Santa Barbara, CA, 93106
USA
almeroth@cs.ucsb.edu

Kimaya Sanzgiri
Dept of Computer Science
University of California
Santa Barbara, CA, 93106
USA
kimaya@cs.ucsb.edu

Abstract

The Internet has evolved from a wired infrastructure to a hybrid of wired and wireless domains. As network access is now provided with much of the last mile being a mobile environment, delivering rich multimedia to mobile users is now a necessity. However, despite the advent of new technology and standards for broadband wireless, there is still an important dilemma over the choice of systems that either achieve high levels of performance or offer easier deployment. The goal of this paper then is to offer insight into this issue by examining the case of media streaming to mobile users through the use of multicast. By specifically considering the debate over network and application layer multicast, we examine a spectrum of possible alternatives and and propose a solution that could be the key to a desired balance between deployment complexity and network performance.

1 Introduction

Delivering rich multimedia to mobile users is no longer regarded as a futuristic network capability. The advent of new technologies, such as broadband wireless access, in addition to ever increasing customer demands, has transformed mobile multimedia into a key differentiator between competitive service providers. In order to meet these new requirements, service and network providers often face a dilemma of adopting a strategy that resides between two extreme points: either reaching the full performance potential by extending the currently deployed network capabilities or to follow a more cost-efficient paradigm where performance and efficiency are compromised in order to protect previous investments. In other words, service providers are faced with the decision of either making do with equipment that is already deployed or upgrading the equipment. Another alternative is sometimes possible: add hardware in a few selected places that operates as a service gateway and performs application layer functionality, e.g. multicast. Frequently related to the principles of the *end-to-end* argument [16] and problems created by placing intelligence in the network, this paper provides insight into this dilemma by focusing on the example of multicast operation for mobile users.

Compared to the one-to-one operation of unicast and the one-to-all of broadcast, *multicast* is a more efficient way of reaching a specific set of network nodes[1]. Implementations of multicast can be broadly realized in two prominent layers: either the network or the application. The first case is most commonly referred to as “IP Multicast” and operates by adding special functionality in the routers of the core network [1]. The second approach is referred to as Application Layer Multicast (ALM) and follows a different paradigm by shifting control to the end hosts themselves [18]. The advantage of IP Multicast is its more efficient use of network resources whereas ALM offers relative deployment simplicity since few or no network modifications are required.

Although the simplicity of ALM was initially associated with important performance penalties, recent advances (like consideration of locality of nodes) have considerably reduced the performance gap [2]. For this reason, ALM has recently attracted much attention in the research community and a number of examples have been the focus of commercial development. It would therefore be reasonable to claim that, for *stationary* nodes, the multicast case provides a good example where the need for deployment simplicity has outweighed the requirement for higher performance efficiency, and therefore, the trend has been more towards adopting ALM-style protocols [7].

However, the introduction of mobility shifts the balance between the two options. In a previous study [8] we have shown that when mobility is introduced, the performance gap between IP multicast and ALM widens considerably. Moreover, the effect of numerous other issues such as network control/management, trust, and node cooperation have had two important implications. First, there is a need to re-evaluate the tradeoffs offered by both ALM and IP Multicast. Second, as numerous extensions have been proposed to the basic schemes, there are now a wide and complicated spectrum of alternative deployment options. Choosing the right balance of complexity and efficiency is a challenging and multi-dimensional problem.

The contribution of this paper is therefore twofold; first we investigate the deployment-versus-performance issue by creating a spectrum and identifying the set of points that represent current and possible solutions. During this investigation we offer both a theoretical evaluation between the choices and a representative set of simulation results. Second, we propose a compromise between the two extremes and offer an initial evaluation. For both aspects of our contribution we are working under the assumption that nodes are operating on a Mobile IPv6 (MIPv6) network. Therefore, although our discussion affects ad hoc scenarios, it is primarily focused on networks that are based on wireless access points.

The remainder of this paper is thus organized as follows. First we explain the motivation for our work by investigating the background and related work. Second we analyze the range of existing alternatives between the two extreme points and we provide a theoretical evaluation for each option. Next we describe a proposed solution followed by qualitative evaluation results. Finally we conclude with a summary of the main points.

2 Background

Current solutions for providing multicast to mobile nodes can be categorized into four groups: (1) (standard) native IP multicast, (2) extended (native) IP multicast, (3) application layer multicast, and (4) extended application layer multicast. This section offers a basic description of each of the four categories and provides a basis for comparison in the next section.

2.1 Standard IP Multicast

The operation of standard IP Multicast can accommodate mobility assuming it operates on a MIPv6 network [13]. MIPv6 is a protocol which allows nodes to remain reachable while moving around in the IPv6 Internet. Mobile nodes may receive packets in one of two ways. In *Reverse Tunneling* a router in the home domain (called the *Home Agent*) receives packets from the original source and tunnels them to the Mobile Node's location. In *Optimized Routing* packets are transmitted from the original source and addressed directly to the new location.

IP Multicast operation can be based on either of these two models. In *Remote Subscription* the node joins the group directly via the multicast router to which the Mobile Node is connected. The disadvantage with this solution is that a rapidly moving node creates a significant *control* load on the network by requiring the tree to be frequently changed. Alternatively, in *Home Subscription* the Mobile Node joins the group via a bi-directional tunnel to its Home Agent. Membership messages are tunneled to the Home Agent, which then joins the group. Data packets that are then received are forwarded to the agent via the tunnel. While the use of Reverse Tunneling can ensure the multicast trees are independent of the Mobile Node's movement, the delay between the source (and then through the Home Agent) and the Mobile Node may be significant. In addition, the delivery tree from the Home Agent to the Mobile Node in such circumstances relies on unicast encapsulation and is therefore bandwidth inefficient compared to native multicast.

2.2 Extended IP Multicast

Although the two basic MIPv6 multicast schemes offer a functional platform, a plethora of limitations have been identified [15]. Typical examples include *tunnel convergence* and *handover* issues. The first occurs in the *Home Subscription* scenario when members of the same group reside in the same domain but are served through different Home Agents. The result is redundant packet transmission since many Home Agents tunnel the same data to the same domain. *Remote Subscription* is affected by handovers as members have to re-join the multicast tree from the new location. This implies that the mobile receiver must not only wait for the next membership query message from the local multicast router but also for the multicast tree to be reconfigured. Even in the case of *Home Subscription*, there are time delays until the node acquires the new care-of address and notifies the Home Agent (which then has to tunnel the incoming packets). As a result, handovers cause various adverse packet effects like delay, loss, jitter and packet duplication.

In order to overcome the various performance issues, several extensions have been proposed. As the various approaches adopt different styles and levels of complexity, the result is a large number of deployment options. The following is a brief sample:

- **Dynamically adaptive systems.** This type of solutions attempt to “adapt” the routing of packets according to the current location of the mobile node. A representative examples of this category is RBMoM [11]. RBMoM is a balance between the Home and Remote Subscription mechanisms. It defines a special role, the

Multicast Home Agent (MHA), which is the node responsible for serving the end users. Initially the operation is based on the Home Subscription scenario as the MHA is the Home Agent itself. It continues to be so as long as the mobile receiver falls within a prespecified hop range. When the receiver goes beyond this range, the role of the MHA is taken by a foreign router, thus switching the operation to Remote Subscription. In general, adaptive solutions such as the RBMoM either propose the insertion of special routers in the core of the network or the complex configuration of existing ones. In both cases deployment complexity is a major concern.

- Hierarchical designs. In the protocols of this style ([12, 19]), there is an attempt to hide micro-mobility of nodes by deploying a hierarchy of routers in the core network. As a result, backbone routers are less frequently concerned about node mobility as this is handled by routers closer to the edge. Although effective, such schemes assume wide scale deployment of specifically configured routers in order to form the hierarchy tree. Deployment considerations are once again a significant burden.

- Proactive schemes. In order to reduce the packet losses caused by handovers, these protocols apply more proactive schemes by trying to guess the next base station. Although efficient in the small scale, these solutions are associated with high complexity and increased traffic in order to be realized. Therefore applying these techniques in the global Internet is a difficult task. Representative examples of this category include Mobicast [17] and MSA [20].

Overall, experimental evaluations have shown that each of the above solutions can address a specific set of problems. However, these gains have to be considered against the additional deployment complexity.

2.3 Standard ALM

ALM is an attempt to overcome the complexity of native multicast by sacrificing a portion of the network efficiency gains for increased deployability. Packets are transmitted through standard unicast messages while replication takes place on the end hosts themselves. Although not as efficient as IP Multicast, ALM has two major advantages. First, as the operation is controlled by the end devices, it manages to eliminate the need for additional support from network routers. Second, it simplifies a number of issues such as congestion control, pricing models and protocol interoperability. In general, ALM protocols take advantage of the combination of protocol flexibility, application compute power, and the relative simplicity and maturity of unicast technology, therefore simplifying the aforementioned issues.

Similar to IP Multicast, ALM protocols can accommodate mobility simply based on the operation of MIPv6. Nodes can move while receiving packets through the standard MIPv6 unicast techniques of *Reverse Tunneling* or *Optimized Routing*. However, as ALM solutions claim independence to the underlying network topology, the introduction of mobility becomes an interesting, and at times, critical factor. Even if we claim that a protocol like MIPv6 handles all low level mobility intricacies, the question becomes whether ALM is still effective.

The general concern arises from the fact that mobility breaks two basic assumptions of ALM protocol designs. First, ALM protocols depend on the relative stability of the network in order to exploit locality information for forming efficient trees. Node mobility breaks this assumption since ALM schemes are often not designed to handle rapid location changes. Second, although robustness has always been an important concern for ALM, mobility again exacerbates the problem. This is because each end node is a link in the forwarding chain of the packets, and the losses of even a single user are propagated down the chain to the rest of the nodes. As a result even stationary nodes will suffer because of packet losses caused by a moving node further up the forwarding chain.

Overall, the more nodes move, the higher the impact on the operation of the ALM protocol [8]. In *Optimized Routing* the tree has to be frequently re-constructed whereas in *Reverse Tunneling* the ALM protocol cannot take advantage of the locality of nodes (since packets always have to go through the Home Agents). The challenge now becomes how to deal effectively with mobility while still having a reasonably simple protocol that is efficient.

2.4 Extended ALM

Since ALM protocols are based on the assumptions of relative network stability and node robustness, we are not aware of any protocols that specifically consider mobile users. Although protocols on ad hoc multicast offer part of the solution, they ignore the key problem of how to transition between wired and wireless domains. Notable exceptions such as our own work [9] and [5] suggest a hybrid model and special proxies located throughout the network. These results show that even if not as efficient as the native multicast solutions, such models offer a better alternative compared to the standard ALM schemes. Nevertheless, the need for special functionality in the core network once again raises the level of deployment complexity.

3 Complexity Versus Performance

In this section we examine the issue of measuring performance efficiency over deployment complexity for the existing solutions described in the last section. The goal is to present the specific characteristics of each alternative in order to both clarify the key advantages and to show the need for additional choices. Overall, each of the four categories satisfies different needs (see Figure 1). At the one extreme, extended IP Multicast offers the best performance but faces the most significant deployment requirements. On the opposite side, a standard ALM solution is a relatively simpler alternative but at the same time the least efficient. The remainder of this section first looks at the performance and then at two aspects of complexity: technical and deployment.

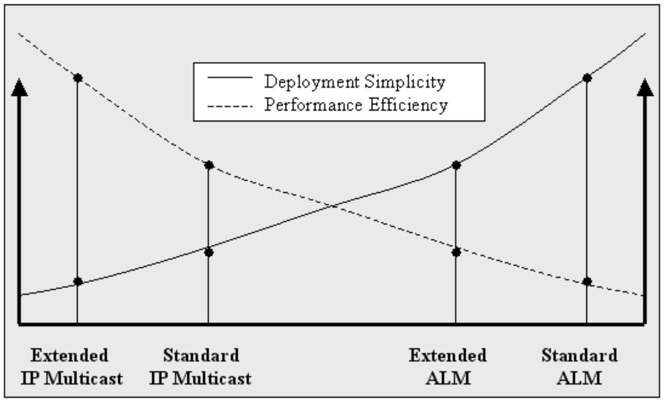

Figure 1. Complexity versus efficiency for current solutions.

3.1 Performance

In a previous study [8] we have shown that, in terms of performance, standard IP Multicast has a clear advantage over ALM. Our comparison was based on the following three metrics:

- **Data throughout.** The ratio of total received packets to those that should have been received assuming no loss.
- **Relative Delay Penalty (RDP).** The ratio of the overlay tree size to the size of the IP multicast tree. The smaller the value, the better for ALM since it means then it more closely matches the performance of IP multicast.
- **Link Stress.** The number of identical packets sent by a node over a particular link. In the case of IP Multicast for stationary nodes, this number is always equal to 1.

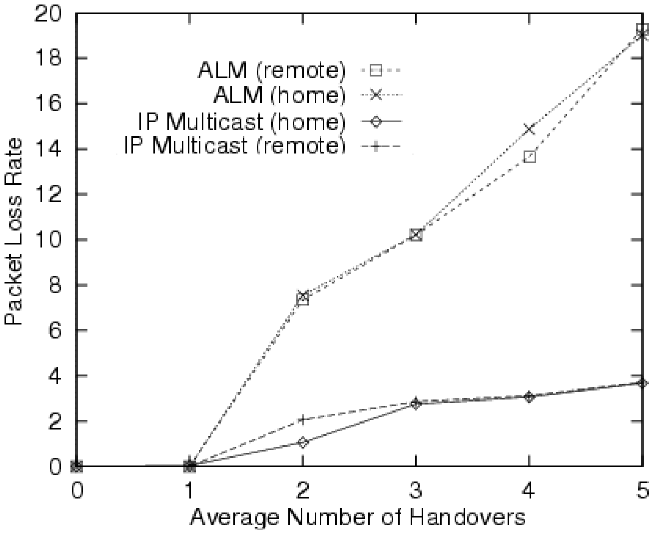

Figure 2. Packet drop rates for ALM and IP Multicast.

Our results show that, in terms of throughput, low mobility gives no major advantage to IP Multicast, i.e. the relative performance of IP Multicast and ALM is essentially the same as the case of no mobility. However, as nodes start to increase their speed, ALM experiences additional packet loss. At its worst, ALM suffers about four times the loss of IP Multicast. This is shown in Figure 2 where the x-axis displays how many handovers occurred on average on each node during the simulation period. The y-axis shows the percentage of lost throughput. This was calculated from the total received packets over those that should have been received. With a number of tests run on 100 mobile nodes (on a network consisting of 200 routers), both IP multicast schemes (*home* and *remote* representing *Home* and *Remote Subscription* respectively), considerably outperform the two possible implementations of the ALM scheme (*home* for *Reverse Tunneling* and *remote* for *Optimized Routing*). The main reason is that with IP Multicast when a node moves, the packet loss is restricted to that specific node whereas in ALM the loss along the overlay path is additive. For example if in an ALM tree Node A is above Node B who is above Node C in the tree and if Node A has an *x%* loss rate due to handovers, Node B would have at best *x%* (plus *y%* due to its own movements) and Node C would have at least (x+y)% loss.

In terms of RDP, with low mobility, IP Multicast performs much better: on the order of four to five times better than ALM. When mobility is high, IP Multicast still performs better but the improvement is less: an RDP ratio of two to one.

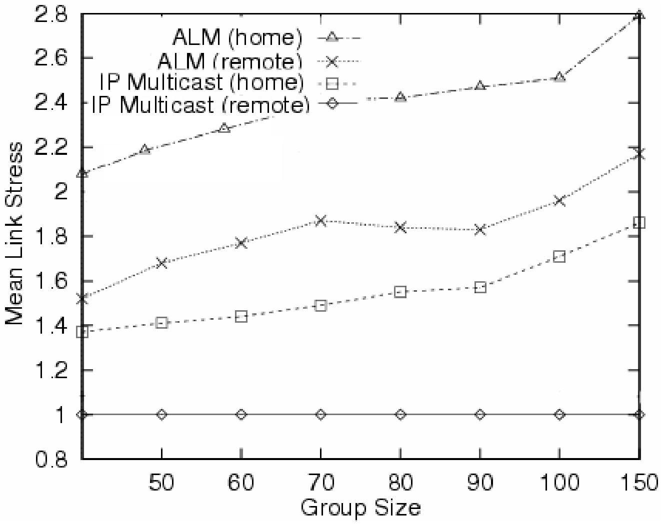

Figure 3. Mean Link Stress for mobile hosts.

Finally, Link Stress is considerably higher for ALM when compared to IP Multicast (around 1.7 times) and increases along with the group size (Figure 3). Overall, ALM suffers both when mobility is low and when it is high. Low mobility gives better robustness but very high RDP. High mobility gives better RDP values but robustness is poor.

3.2 Complexity

In this part we consider two kinds of complexity: protocol overhead and deployment considerations. Each of the two types is discussed below.

Protocol complexity can be measured by both the number of control messages and the level of intelligence that is required by the involved entities. Despite its efficiency, IP Multicast is associated with various complexities, including the following:

- **Inter-domain operation.** Due to the plethora of proposed protocols, interoperability among the various domains is a serious concern. Solutions such as the Border Gateway Multicast Protocol (BGMP)[10] are regarded to be too complicated and expensive to be implemented.

- **Address allocation.** Since the current multicast address space is unregulated, special precautions have to be taken in order to avoid clash of different groups that have the same address. Again, existing proposals, such as the Multicast Address-Set Claim (MASC) protocol[10], are not satisfactory for long term solutions.

- **Miscellaneous technical issues.** Further technical characteristics, such as the provision of QoS, security and billing functionality, have long been regarded as complicated problems[6].

Mobility introduces further considerations as described below:

- There has to be a critical decision about the choice between Home and Remote Subscription. Home Subscription is simpler and offers handover transparency to multicast operation since the tree remains the same (only the Home Agent has to change the tunnel destination). However it suffers from *triangular routing* (packets have to go through the Home Agent), overhead due to the encapsulation and decapsulation of the packets, and *tunnel convergence*. Remote Subscription is regarded as more efficient but results in frequent reconstruction of the branches of the multicast tree. Moreover it assumes that the same IP Multicast protocol will be deployed in all visited domains.

- End nodes have to be capable of handling additional computation complexity. For example, packet duplication can occur during a handover since the node may receive the same data from two neighbor domains (assuming they have both joined the same multicast group). Also, if the receiver moves out of the multicast scope (maximum hop count allowed), data will not be received.

The many deployment complexities of IP multicast were the key reasons for the creation of ALM protocols. Interdomain communication and address allocation are now less of an issue while the unicast technologies on QoS and billing are more robust and mature. Nevertheless, as already explained in the previous section, ALM protocols suffer considerably more than IP Multicast when mobility is introduced. As scalable ALM solutions are often associated with more complex tree building processes, continued readjustment of overlay trees raises the protocol complexity considerably. Overall, in terms of protocol complexity, although ALM is better positioned than IP Multicast, mobility implies that we are in need of another alternative.

In terms of deployment considerations, the following two are the most important:

- **Investment costs.** We regard investment costs to be influenced by two important aspects: first, the possibility of updating or replacing nodes in the core network, such as routers, and second, the potential to configure end devices. Configuring core routers is not only a complicated process but also has to face the hesitation of ISPs to actually perform this task. In addition, as backbone routers are usually heavily loaded nodes with very high efficiency requirements, particular emphasis is given to making them as simple and as efficient as possible. For this reason, it is generally preferable to push the complexity towards the edge of the network[16]. However, configuring end devices may prove to be equally problematic since it requires the users to be proactive, install software, and generally participate in any scheme dictated by the service or network providers.

- **Domain independence.** The lack of global coordination across the various ISPs has resulted in a mosaic of different administrative domains in the Internet. As a result, there is a high degree of heterogeneity not only in terms of technical capabilities but also in terms of commercial interests. Although cross-domain operation is realized through Service Level Agreements (SLAs), the interoperability of underlying protocol communication and technical capabilities cannot be assumed. This is another very important reason why standard IP Multicast has not been widely deployed, since the large number of protocol specifications resulted in multiple “multicast islands”[6]. Communication between these islands compromises the efficiency gains of multicast and raises the deployment costs (gateways and tunneling of packets). An ALM solution offers a better alternative since it only uses standard unicast messages, a type of communication which is simpler and more efficiently handled across the various domains.

Overall, native multicast is hard to deploy because it requires changes to the complete spectrum of network devices: backbone routers, edge routers, switches and hosts. ALM is an attempt to try and limit the number of places where changes are required by pushing intelligence to the edge. But requiring complexity in the edge nodes has the additional requirement of getting users to cooperate and install software components. So, we advocate a solution that ”concentrates” required changes. No changes are required in the core and as few changes as possible are required in end hosts.

4 Proposed Solution

This section presents the Intelligent Gateway Multicast (IGM) protocol and consists of two parts. First we describe the main considerations for our design and then we describe our proposed architecture.

4.1 Design Considerations

The aim of our desired solution is to achieve two goals; first, to achieve satisfactory levels of network performance and second, to avoid the limitations of a complex solution requiring deployment in all components of an end-to-end path.

- **Avoid complexity at the backbone routers.** Although we would like to make use of IP Multicast, we cannot assume its existence. If IP Multicast is not available, we do assume an ALM protocol or some mechanism to provide one-to-many communication.

- **Need for operation control.** One general disadvantage of ALM protocols that is applicable in this situation is that in ALM the end nodes that participate must be trusted to behave correctly. We believe this assumption should be avoided as much as possible. Because our perspective is one of actually trying to deploy a real-world system, we believe that an ISP will want some level of operation control, i.e. an ISP will want to control how multicast is provided and to whom[14]. For example, an ISP might not want to provide a streaming data service to a rapidly moving node that is often in the process of changing access points and constantly requiring a change in a multicast tree to which it belongs.

- **Minimize the impact of mobility.** User mobility, as just mentioned, can place significant performance overhead on core and edge network elements. Apart from the performance issues in terms of the network savings[8], the penalty of expecting routers to (nearly) continuously modify the tree is a very expensive proposition for both ALM and native multicast. Therefore there is a need to hide as much of the effects of node movement as possible.

- **Provide service differentiation between nodes.** Device heterogeneity is a well known issue and has attracted a significant amount of attention by the research community. There is a need to treat nodes differently based not only on their capabilities, but also on other factors like the characteristics of their node movement. The need to treat nodes differently depending on their speed is an issue we have shown previously to have a significant impact on performance[9].

Considering all these four issues, we propose a solution that is based on multicast support in *intelligent gateways*. These devices are not new routers, but intelligent machines that are co-located with the radio access stations. The complexity of providing a service like multicast is concentrated in these devices. Although the installment of such machines is associated with a certain investment cost, we expect this to be considerably reduced compared to the configuration or replacement of core routers. The details of IGM are described below.

4.2 Proposed Architecture

The core element of our architecture is the *gateway*, an intelligent node co-located with the radio transmission antenna of any access network (Figure 4). Each *gateway* is then responsible for a number of operations:

- Advertise their presence and set of services to the local mobile nodes[3]. On reception of such advertisements the nodes can then direct *join* messages to the gateway.
- Keep a record of the currently served nodes. Such records will be simple mappings of the mobile node ID and the corresponding Home Agent address.
- Relay packets destined to the mobile nodes as these arrive from the Home Agent of the corresponding node.
- If requested, relay incoming packets to neighboring gateways. This is an attempt to hide the mobility of the node from the Home Agent by relaying packets instead of reconfiguring some of the path. Relaying is described later.
- Act as a firewall and access control point for authorizing user join requests and filtering unwanted traffic[14].

The only requirement for mobile nodes is to keep a history of the most recently used gateways. The importance of this information is explained in the next section.

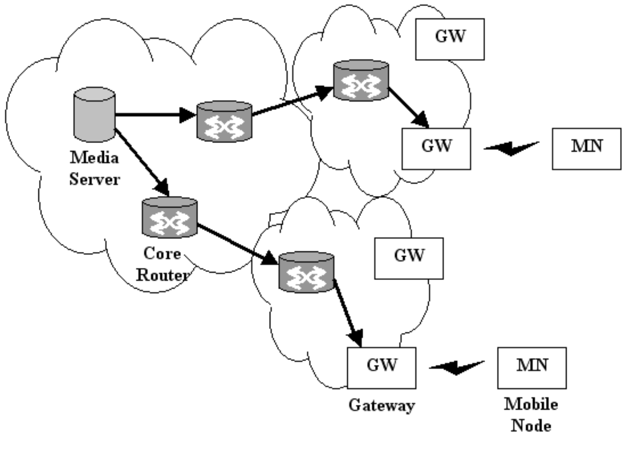

Figure 4. Proposed IGM Architecture.

The architecture does not assume the wide deployment of either native or overlay multicast. The only requirement is that the node will join any available tree through its Home Agent, thus following the *home subscription* mode of the MIPv6 protocol. Although utilization of some kind of one-to-many distribution tree is encouraged, this is not actually a requirement. In the worse case, one-to-many communication can be provided by the content source as a replicated set of unicast streams. The key objective of our protocol therefore is to minimize, as much as possible, the impact of mobility on whatever distribution mechanism is used. Even though this idea has already been suggested by several hierarchical schemes [12, 17], the main difference to our system is that our architecture does not require installation of proxies in the core network but at the edge, on the border between the wired and the wireless domains.

4.3 Protocol Operation

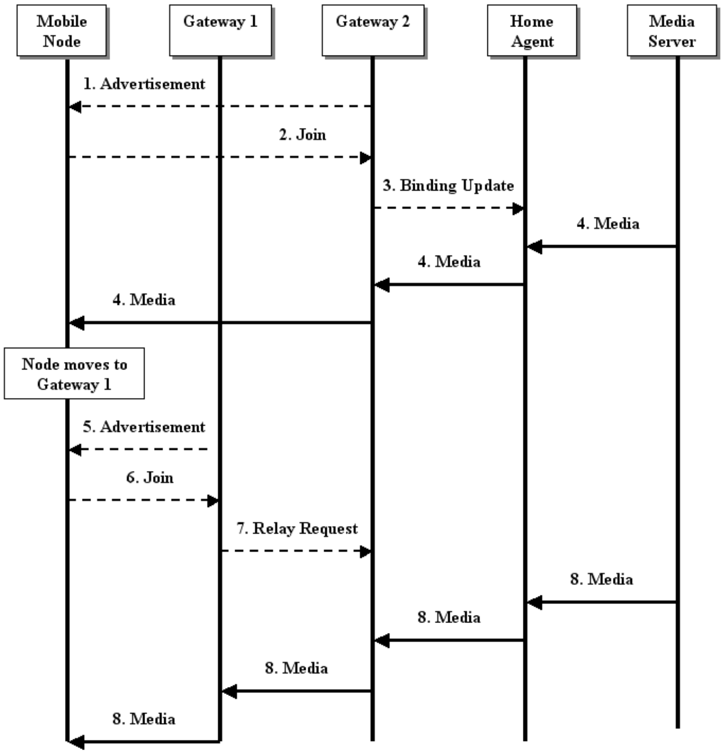

Figure 5. IGM Protocol Messages.

Figure 5 displays the message exchange for our IGM protocol. We also explain each of the steps below.

1. When the node enters a new area, it first discovers the associated gateway. It can do this by (1) including the information as part of the handover, (2) broadcasting a service advertisement, or (3) waiting for a periodic broadcast from the gateway. The most efficient of these choices is for the gateway information to be part of the handover process. In Figure 5, we show an example of a mobile node moving to *Gateway 2*. As part of this communication, the mobile node will discover the gateway's IP address and will receive a

list of available services (e.g. QoS provisions). More generally, we believe there will exist a need for intelligent handovers when possibly many wireless companies are offering varied and competing services[4]. In this scenario, a user device will have the option to select among the offered set of gateways and their differing sets of services.

2. After selecting a gateway, the Mobile Node joins the multicast group by sending a *Join* message. This *Join* must contain the IP address of the last node from which it was receiving multicast content. This address could be another gateway, the mobile node's home agent, or the content source itself. If this is the node's first attempt to join the group, the field for this information will be empty.

3. The Gateway enters a mapping of the mobile node ID and the corresponding *Last Relay* Address. Then it sends a MIPv6 specified *Binding Update* message to the Last Relay. This contains the new address (care-of address) of the Mobile Node to which the Home Agent must tunnel incoming packets. Based on the operation of this protocol, the care-of address corresponds to the IP address of the this new gateway.

4. As media content arrives to the Last Relay node (either through IP Multicast, ALM, or a tunnel), the Last Relay tunnels it to the recorded care-of address. As a result, the encapsulated packets are sent to Gateway 2 (see Figure 3). This gateway then checks the cache of Mobile Node to Last Relay mappings, and relays the content to the corresponding mobile node.

5. If at a later stage the node moves to another area and wants to associate with a new gateway (e.g. Gateway 1), it then associates with the new gateway and repeats the process.

6. Similar to before, after associating with the new gateway, the mobile node will send a *Join* message. This time, the Last Relay address will be for Gateway 2.

7. At this stage, Gateway 1 sends a *Relay Request* message to Gateway 2.

8. Gateway 2 now receives traffic from its own Last Relay, forwards it to Gateway 1 who then relays it to the mobile node. Relay links are kept alive based on a soft state mechanism where links are dropped if a keep-alive message has not been received within a specified time frame.

Beyond this exchange of information, there are three additional mechanisms used to improve the efficiency of the system.

- In order to avoid long forwarding chains between the involved gateways, we propose a “path check” mechanism for gateways. After processing a join message and starting the process of forwarding data packets, the gateway sends a ping message to the source of the packets. By comparing the hop count (or other quality metrics) of the ping response to the hop count of relayed data packets, the gateway can decide whether it would be more efficient to join the group directly. In other words, if the gateway believes that the relay path is too inefficient, it can initiate a join to the source directly. The threshold for efficiency can be determined in a variety of ways according to local circumstances. Here we have simply described it as a hop count computation.

- Another potential limitation of the basic scheme is that a node may continually move between two or three gateways (the ping-pong effect). As a result, even though the hop count to the source may by smaller than that of the relay path, it would clearly be desirable to prevent tunneling loops or frequent attempts to re-configure the path. In order to avoid such scenarios, upon receiving a new join message, a gateway will check to verify that it has not recently been a relay node for the mobile node. If this is in fact the case, the gateway will reject the relay request.

- Finally, as suggested by the goal of the whole protocol, to protect the infrastructure from rapidly moving nodes that generate frequent handovers, there is a need to identify these nodes and provide different operation. To be more specific, even if a gateway discovers that the hop count to the Home Agent is lower than the one from the relay chain, it should wait for a short period of time before making the transition. The reason is that if the node is moving quickly, there will be little use in contacting the Home Agent and re-building the tree. Consequently, if the mobile node informs the Gateway of its speed of movement, the Gateway can decide how to act.

5 Evaluation Framework

Our future work for IGM will be a full evaluation of the protocol in order to quantitatively assess the advantages and disadvantages. Here, we therefore only present a qualitative evaluation and we provide the details of our experimental framework.

5.1 Qualitative Evaluation

In this part we examine how we have addressed the architectural considerations described in Section 4.1.

- **Avoid complexity at the backbone routers.** This is achieved in two ways. First, by being independent of IP Multicast (ALM operation is assumed if this is not present), we exploit existing deployments. Second, although we advocate the installation of specific gateways, we avoid putting the complexity in the core of the network. In general, we anticipate that use intelligent gateways in the borders between wired and wireless domains to become an emerging trend [3, 4].

- **Need for operation control.** Since we believe that an ISP will want some level of operation control, this can now be achieved through the gateways. Even if an ALM-style solution is widely deployed, ISPs can still have some level of control since an important part of the protocol operation will go through their deployed equipment. Filtering and throughput control can be used to allow or impede data delivery.

- **Minimize the impact of mobility.** Since the intelligence is now placed on the gateways, core routers will no longer have to modify the tree in either ALM or native multicast. Although not completely hidden, gateways are effective in abstracting much of the effect of node movement.

- **Provide service differentiation between nodes.** Having the intelligence in the gateways enables one more solution. The need to treat nodes differently depending on their speed is now an issue that can be dealt with in a variety of ways. As the nodes provide their speed or time between handovers to the gateway, the latter can independently decide on the required action. Different ISPs may need to take different courses of action, limiting the access of some and completely blocking the access of others.

5.2 Evaluation Framework

We now describe the framework details for our evaluation. We are running the simulations using a packet level discrete-event simulator written in Java. Topologies for the wired part of the network are power-law graphs consisting of a two-level node hierarchy. Core nodes represent the routers whereas leaf nodes correspond to intelligent gateways. The simulator has been enhanced with basic implementations of MIPv6, IP Multicast (PIM-SM protocol) and a generic ALM protocol. For the latter, in order to capture the most important aspects, we compute a shortest path tree over the complete set of overlay nodes.

The most important parameters in our simulations are as follows:

- **Number of nodes.** 500 routers and gateways

- **Number of receivers.** These are the multicast group members. They fall within a range of 10 to 200.

- **Ratio of mobile receivers.** The percentage of nodes that are mobile. We investigate percentages of 25/

- **Mobility speed.** The average number of handovers (moving from one gateway to the next) in a specific simulation period.

- **Mobility pattern.** The model by which nodes move, their direction and speed. For our simulations we use *Random Waypoint*.

Based on this setup, we are performing our simulations using the following metrics:

- **Data throughput.** The ratio of total received packets to those that should have been received assuming no loss.

- **Relative Delay Penalty (RDP).** The ratio of the overlay tree size to the size of the IP multicast tree. The smaller the value, the better for ALM since it means then it more closely matches the performance of IP multicast.

- **Link Stress.** The number of identical packets sent by a node over a particular link. In the case of IP Multicast for stationary nodes, this number is always equal to 1.

What we expect to find is that for slow moving or stationary nodes, performance will essentially be the same as a comparison between IP multicast and an ALM protocol that does not consider mobility. The case of faster moving nodes will be the more interesting case. What we expect to see is long chains of gateways acting as relays. These chains could be inefficient, but inefficient chains should be eliminated by a gateway who decides to break the chain and join to the original content source directly. The chains that are left should therefore not be too much worse than a more direct path. This path inefficiency though will be countered by many fewer changes in the tree topology (maintained either through IP multicast or an ALM protocol).

6 Conclusions

Compared to the one-to-one operation of unicast and the one-to-all of broadcast, *multicast* is a more efficient way of reaching a specific set of network nodes. Given the media-rich applications now being introduced, efficient delivery of content is a critical service. Implementations of multicast can be broadly realized in two prominent layers: either the network or the application. Although each approach is associated with specific advantages (performance efficiency for IP Multicast and deployment simplicity for ALM), the introduction of mobility introduces several complexities.

In a previous study [8] we have shown that when mobility is introduced, the performance gap between IP multicast and ALM widens considerably. Moreover, the effect of numerous other issues such as network control/management, trust, and node cooperation have two important implications. First, there is a need to re-evaluate the tradeoffs offered by both ALM and IP Multicast. Second, as numerous extensions have been proposed to the basic schemes, we are now faced with a wide and complicated spectrum of alternative deployment options. Choosing the right balance of complexity and efficiency is a challenging and multi-dimensional problem.

In this paper, we have focused on two issues. First, we have investigated the deployment-versus-performance issue by creating a spectrum and identifying the set of points that represent current and possible solutions. Second, we have proposed a compromise between the two extremes with the use of intelligent gateways. We believe that through the use of gateways, which we anticipate will be widely used in the near future for a variety of services, we can concentrate the complexity and additional functionality required to support multicast. Our evaluation in this paper has simply been a qualitative evaluation of the idea and the setup for further experiments. The simulation and evaluation of our scheme is left to future work.

References

[1] K. Almeroth. The evolution of multicast: From the MBone to inter-domain multicast to Internet2 deployment. *IEEE Network*, 14(1):10–20, January/February 2000.

[2] M. Castro, P. Druschel, Y. Hu, and A. Rowston. Exploiting network proximity in peer-to-peer overlay networks. In *International Workshop on Future Directions in Distributed Computing*, Bertinoro, Italy, June 2002.

[3] R. Chalmers and K. Almeroth. A mobility gateway for small-device networks. In *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, Orlando, FL, March 2004.

[4] R. Chalmers, G. Krishnamurthi, and K. Almeroth. Enabling intelligent handovers in heterogeneous networks. Technical report, University of California, Santa Barbara, October 2003. (submitted).

[5] J. Chennikara, W. Chen, A. Dutta, and O. Altintas. Application layer multicast for mobile users in diverse networks. In *IEEE Globecom*, Taipei, Taiwan, November 2002.

[6] C. Diot, B. Levine, B. Lyles, H. Kassem, and D. Balensiefen. Deployment issues for the IP multicast service and architecture. *IEEE Network*, 14(1):78–88, February 2000.

[7] S. Fahmy and M. Kwon. Characterizing overlay multicast networks. In *IEEE International Conference on Network Protocols (ICNP)*, Atlanta, Georgia, November 2003.

[8] A. Garyfalos, K. Almeroth, and J. Finney. A comparison of network and application layer multicast for mobile IPv6 networks. In *ACM Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, San Diego, CA, September 2003.

[9] A. Garyfalos, K. Almeroth, and J. Finney. A hybrid of network and application layer multicast for mobile IPv6 networks. In *Workshop on Large-Scale Group Communication (LSGC)*, Florence, Italy, October 2003.

[10] S. Kumar, P. Radoslavov, D. Thaler, C. Alaettinoglu, D. Estrin, and M. Handley. The MASC/BGMP architecture for inter-domain multicast routing. In *ACM Sigcomm*, Vancouver, Canada, August 1998.

[11] C. Lin and K. Wan. Mobile multicast support in IP networks. In *IEEE Infocom*, Tel-Aviv, Israel, March 2000.

[12] H. Omar, T. Saadawi, and M. Lee. Multicast support for mobile-IP with the hierarchical local registration approach. In *ACM International Workshop on Wireless Mobile Multimedia (WOWMOM)*, Boston, Massachusetts, August 2000.

[13] C. Perkins and D. Johnson. Mobility support in ipv6. In *ACM Mobicom*, Rye, New York, November 1996.

[14] K. Ramachandran and K. Almeroth. MAFIA: A multicast management solution for access control and traffic filtering. In *IEEE/IFIP Conference on Management of Multimedia Networks and Services (MMNS)*, Belfast, Northern Ireland, September 2003.

[15] I. Romdhani, et. al. IP mobile multicast: Challenges and solutions. *IEEE Communications Surveys and Tutorials*, 6(1), First Quarter 2004.

[16] J. Saltzer, D. Reed, and D. Clark. End-to-end arguments in system design. *ACM Transactions in Computer Systems*, 2(4), November 1984.

[17] C. Tan and S. Pink. Mobicast: A multicast scheme for wireless networks. *Kluwer Mobile Networks and Applications*, 5(4), 2000.

[18] S.-W. Tan, G. Waters, and J. Crawford. A survey and performance evaluation of scalable tree-based application layer multicast protocols. Technical Report 9-03, University of Kent, July 2003.

[19] Y. Wang and W. Chen. Supporting IP multicast for mobile hosts. *Kluwer Mobile Networks and Applications*, 6(1), February 2001.

[20] J. Wu and G. Maguire Jr. Agent based seamless IP multicast receiver handover. In *IFIP Personal Wireless Communications (PWC)*, Gdansk, Poland, September 2000.